Tunnelling density of states at Coulomb blockade peaks

Nicholas Sedlmayr, Igor V. Yurkevich and Igor V. Lerner
School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, United Kingdom
(Dated: 25 July 2006)

We calculate the tunnelling density of states (TDoS) for a quantum dot in the Coulomb blockade regime, using a functional integral representation with allowing correctly for the charge quantisation. We show that in addition to the well-known gap in the TDoS in the Coulomb-blockade valleys, there is a suppression of the TDoS at the peaks. We show that such a suppression is necessary in order to get the correct result for the peak of the differential conductance through an almost close quantum dot.

I. INTRODUCTION

The Coulomb blockade in quantum dots is one of the most thoroughly investigated and best understood phenomena in modern condensed matter physics (for reviews see[1,2]). Starting from the earliest papers on the subject,[3,4,5,6], the attention was focused on the conductance through Coulomb-blockaded dots and on the corresponding statistics of heights, widths, and positions of the Coulomb-blockade peaks in conductance. A complimentary approach was to describe the tunnelling density of states (TDoS) in the dot[7,8,9]. While also leading to the description of conductance and its distribution it was of considerable interest by itself. The reason is that a singularity in the TDoS in the zero-dimensional dot potentially bridges the zero-bias anomaly in the TDoS of interacting electrons in the metallic regime[10] with the Coulomb gap in the insulating regime[11].

Qualitatively, behaviour of the TDoS found in[7,8,9] has a clear origin. The Coulomb blockade makes the energetic cost of putting an extra electron in the dot of the order of the charging energy, $E_{\rm c} = e^2/C$ (C is an effective capacitance). Consequently, it leads to a low-energy gap in the TDoS stretching up to $E_{\rm c}$ at low temperatures, $T \ll E_{\rm c}$. By tuning the gate voltage applied to the dot one periodically reaches degeneracy points where the energy of having N and $N+1$ electrons in the dot is the same. At these points, i.e. at the peaks in the Coulomb blockade regime, tunnelling is no longer suppressed and one expects the gap in the TDoS to close. In this Letter we show that, although the TDoS at the peaks remains finite, its zero-energy value is of half of its value at $\varepsilon \gtrsim E_{\rm c}$. We then demonstrate that such a suppression is necessary to restore the correct results[3,4,5,6] for the peak conductance.

We will use technique similar to that introduced by Kamenev and Gefen[8] in their original calculation of the TDoS in the quantum dot, representing the electron Green's functions as functional integrals, albeit we use the Keldysh representation in the form developed in[12] rather than the Matsubara one as in[8]. The main difference in our approach is that in the saddle-point approximation, we take into account all the electron winding numbers. This is necessary to get the correct expression for the TDoS at the Coulomb blockade peaks and also is the only consistent way (in the discussed context) to get the electron number quantization both at the peaks and in the valleys. Naturally, this could be obtained either in Keldysh or in Matsubara formalism. The necessity of including all the winding numbers has already been stressed in[13] for the problem of a granular medium comprised of an array of quantum dots in the Coulomb blockade regime. However, the effective averaging over chemical potentials of the dots inevitable for the granular system, gives in this case only the value of the TDoS in the valleys.

We consider the problem in the ergodic (zero-dimensional) limit. It is straightforward in our formalism to include a correction from spatially inhomogeneous modes; however, this correction is small as $1/g$ (where $g \sim 1/\tau_{\rm dw}\delta \gg 1$ is the dimensionless conductance of the dot and $\tau_{\rm dw}$ is the electron's mean dwelling time in the dot) and will not differ from that found previously[8]. Thus we start with the standard 'universal' Hamiltonian for a zero-dimensional system[2,8,14], keeping there only the charging term:

$$\hat{H} = \hat{H}_0 + \frac{E_{\rm c}}{2}\left(\hat{N} - N_g\right)^2 . \tag{1}$$

$\hat{H}_0$ is the Hamiltonian of free electrons in a random potential V, $\hat{N}$ is the electron number operator, eN_g is the neutralising background charge (governed by the gate voltage for the standard quantum dot). It is convenient to represent the appropriate Green's function in terms of the functional integral in the Keldysh technique[12],

$$iG(\mathbf{r},t;\mathbf{r}',t') = Z^{-1}\int \mathcal{D}\bar{\psi}\mathcal{D}\psi\, \psi(\mathbf{r},t)\bar{\psi}(\mathbf{r}',t')\, {\rm e}^{iS[\psi]}\,, \tag{2}$$

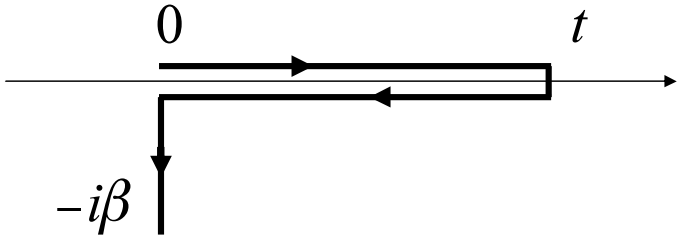

FIG. 1: The 'interaction' Keldysh contour[16].

where $Z = \int \mathcal{D}\bar{\psi}\mathcal{D}\psi\, \mathrm{e}^{iS[\psi]}$ and the action $S[\psi] = S_0[\psi] + S_{\rm c}[\psi]$ is given by

$$\begin{aligned} S_0[\psi] &= \int_K \mathrm{d}t \int \mathrm{d}\mathbf{r}\, \bar{\psi}(\mathbf{r},t)\left[i\partial_t - \hat{\xi}\right]\psi(\mathbf{r},t)\,, \qquad \hat{\xi} \equiv \frac{\hat{p}^2}{2m} + V - \widetilde{\mu} \\ S_{\rm c}[\psi] &= -\frac{E_{\rm c}}{2}\int_K \mathrm{d}t\, N^2(t), \qquad N(t) = \int \mathrm{d}\mathbf{r}\, \bar{\psi}(\mathbf{r},t)\psi(\mathbf{r},t)\,. \end{aligned} \tag{3}$$

The time integrals are taken along the standard 'interaction' contour[15,16], Fig.1, and $\widetilde{\mu} = \mu + E_{\rm c}N_g$ on the thermodynamic (imaginary) part of the contour and $\widetilde{\mu} = E_{\rm c}N_g$ on its dynamical (real) part. This corresponds to a Green's function defined as an average with the grand canonical Gibbs density matrix; its time evolution is described in the Heisenberg representation with a Hamiltonian, Eq. (1), that does not contain μ. In the zero-dimensional regime considered here it is convenient to expand Grassmann fields $\psi(\mathbf{r},t)$ and $\bar{\psi}(\mathbf{r},t)$ in terms of free electron eigenfunctions, introducing fields $c(t)$ and $\bar{c}(t)$ which depend on time only:

$$\psi(\mathbf{r},t) = \sum_n \psi_n(\mathbf{r})\, c_n(t), \quad \hat{\xi}\psi_n(\mathbf{r}) = \xi_n\, \psi_n(\mathbf{r}), \quad \xi_n = \varepsilon_n - \widetilde{\mu}.$$

The Green's function becomes

$$G(\mathbf{r},t;\mathbf{r}',t') = \sum_n \psi_n(\mathbf{r})\bar{\psi}_n(\mathbf{r}')\, G_n(t,t') \tag{4}$$

In the zero-dimensional system under consideration positions $\mathbf{r}$ and $\mathbf{r}'$ in Eq. (4) are indistinguishable so that observable quantities can be found from

$$G(t,t') \equiv \int \mathrm{d}^d r\, G(\mathbf{r},t;\mathbf{r},t') = \sum_n G_n(t,t')\,. \tag{5}$$

Decoupling the charging term $S_{\rm c}$ in Eq. (3) with the help of the standard Hubbard-Stratonovich transformation leads to (i) replacing $S_{\rm c}$ by the bosonic action

$$S[\phi] = -\frac{1}{2E_{\rm c}}\int_K \mathrm{d}t\, \phi^2(t)\,, \tag{6}$$

and (ii) substituting $i\partial_t - \phi$ for $i\partial_t$ in the action S_0. To calculate the functional integral over the fermionic fields $c_n(t)$ and $\bar{c}_n(t)$, and thus $G(t,t')$, we notice that for any function $\varphi(t)$

$$\begin{aligned} Z_0^{-1}\int \mathcal{D}\bar{c}_n\, \mathcal{D}c_n\, \mathrm{e}^{i\int_K \mathrm{d}t\, \bar{c}_n[i\partial_t + \varphi(t)]c_n} &= 1 + \mathrm{e}^{i\int_K \mathrm{d}t\, \varphi(t)}, \\ Z_0^{-1}\int \mathcal{D}\bar{c}_n\, \mathcal{D}c_n\, c_n(t)\bar{c}_n(t')\, \mathrm{e}^{i\int_K \mathrm{d}t\, \bar{c}_n[i\partial_t + \varphi(t)]c_n} &= \mathrm{sgn}(t,t')\, \mathrm{e}^{i\int_{t'}^{t}\mathrm{d}\tau\, \varphi(\tau)}\,. \end{aligned}$$

The time ordering here is along the contour in Fig. 1 so that the integral in the second line above is also taken along the contour and $\mathrm{sgn}(t,t')$ equals 1 (or -1) when t precedes (or goes after) t' on the contour. The bosonic field is not included in the normalization factor Z_0 in the above expressions.

Note the analogy with the calculations in the Matsubara technique in Ref.[8]: the bosonic field can be gauged out by a shift in the fermionic field but for the initial conditions in the imaginary time integral. These do not allow one

to get rid in this way of the zero-frequency field component. In our calculations ϕ_0 is the precise analogue of this component. Hence

$$G_n(t,t') = \frac{-i\,\mathrm{sgn}(t,t')}{\mathcal{Z}} \int \mathcal{D}\phi\, \mathrm{e}^{iS[\phi]}\, \Xi_n(\phi_0)\, \mathrm{e}^{\int_{t'}^{t} \mathrm{d}\tau[-i\xi_n+\phi(\tau)]}\,, \qquad \mathcal{Z} = \int \mathcal{D}\phi\, \mathrm{e}^{iS[\phi]}\, \Xi(\phi_0)\,. \tag{7}$$

Here we have defined the grand canonical partition function, $\Xi(\phi_0)$, and the grand canonical partition function with the n^{th} level excluded, $\Xi_n(\phi_0)$, with energy levels in both shifted by the charging effects:

$$\Xi(\phi_0) \equiv \prod_m \left(1+\mathrm{e}^{-\beta\xi_m+\phi_0}\right)\,, \qquad \Xi_n(\phi_0) \equiv \Xi(\phi_0)\left(1+\mathrm{e}^{-\beta\xi_n+\phi_0}\right)^{-1}\,. \tag{8}$$

It is convenient to expand $\Xi(\phi_0)$ and $\Xi_n(\phi_0)$ in Eq. (8) in terms of the canonical partition functions:

$$\begin{aligned} \Xi(\phi_0) &= \sum_{N=0}^{\infty} Z_N\, \mathrm{e}^{(\beta\tilde{\mu}+\phi_0)N}\,, \qquad Z_N = \oint \frac{\mathrm{d}\theta}{2\pi}\, \mathrm{e}^{-iN\theta} \prod_m \left(1+\mathrm{e}^{-\beta\varepsilon_m+i\theta}\right)\,, \\ \Xi_n(\phi_0) &= \sum_{N=0}^{\infty} Z_N(\varepsilon_n)\, \mathrm{e}^{(\beta\tilde{\mu}+\phi_0)N}\,, \quad Z_N(\varepsilon_n) = \oint \frac{\mathrm{d}\theta}{2\pi}\, \mathrm{e}^{-iN\theta} \prod_{m\neq n} \left(1+\mathrm{e}^{-\beta\varepsilon_m+i\theta}\right)\,. \end{aligned} \tag{9}$$

The meaning of $Z_N(\varepsilon_n)$ can be seen from its formal definition:

$$\frac{Z_N(\varepsilon_n)}{Z_N} = \frac{\mathrm{Tr}_N\left(c_n c_n^{\dagger} \mathrm{e}^{-\beta\hat{H}_0}\right)}{\mathrm{Tr}_N\left(\mathrm{e}^{-\beta\hat{H}_0}\right)} = 1 - F_N(\varepsilon_n)\,, \tag{10}$$

where $F_N(\varepsilon_n)$ is the distribution function in the system of N noninteracting electrons, as the charging energy in the Hamiltonian (1) is just a constant when the number of electrons N is fixed. Note that in order to get a consistent description of the charge quantization in the quantum dot, it is crucial to have all winding numbers m in the expansion (9).

On substituting this expansion and a similar one for Ξ_n into Eq. (7), and performing the Gaussian integration over the fields ϕ, we find all the Keldysh components of the Green's function. Thus, the Fourier transform of $G^{>}(t,t')$ (where t is on the lower, and t' is on the upper part of the Keldysh contour, Fig.1) is given by:

$$G^{>}(\varepsilon) = -\frac{2\pi i}{\mathcal{Z}} \sum_n \sum_{N=0}^{\infty} e^{-\beta E_N}\, Z_N(\varepsilon-\Omega_N)\, \delta(\varepsilon-\varepsilon_n-\Omega_N)\,, \tag{11}$$

where

$$\mathcal{Z} = \sum_{N=0}^{\infty} \mathrm{e}^{-\beta E_N}\, Z_N\,, \qquad E_N \equiv \frac{E_{\mathrm{c}}}{2}\,(N-N_g)^2 - \mu\, N\,, \qquad \Omega_N \equiv E_{\mathrm{c}}\left(N+\tfrac{1}{2}-N_g\right).$$

Now we can effectively average over disorder simply by substituting the mean TDoS of noninteracting electrons, ν_0, for $\sum_n \delta(\varepsilon-\varepsilon_n-\Omega_N)$, with the assumption that the TDoS is smooth in any realisation of disorder which is valid when the mean (or typical) level spacing δ is much smaller than T. Next we use Eq. (10) to obtain

$$G^{>}(\varepsilon) = -\frac{2\pi i\nu_0}{\mathcal{Z}} \sum_N \mathrm{e}^{-\beta E_N}\left[1-F_N(\varepsilon-\Omega_N)\right] \approx -\frac{2\pi i\nu_0}{\mathcal{Z}} \sum_N \mathrm{e}^{-\beta E_N}\left[1-f(\varepsilon-\Omega_N)\right], \tag{12}$$

where we have used that $F_N(\varepsilon-\Omega_N)$ for $N \gg 1$ is approximately the same as the (grand canonical) Fermi-Dirac distribution function $f(\varepsilon-\Omega_N)$ with the 'chemical potential' μ_0 of order $N\delta$ (which is small as compared to Ω_N and thus discarded). Note that this obvious relation between canonical and grand canonical expressions can be easily confirmed directly by using the definition (10) and calculating Z_N and $Z_N(\varepsilon_n)$, Eq. (9), in the saddle point approximation.

The TDoS can be found from the standard formula

$$\nu(\varepsilon) = \frac{i}{2\pi}[G^R(\varepsilon) - G^A(\varepsilon)] = \frac{i}{2\pi}[G^{>}(\varepsilon) - G^{<}(\varepsilon)]\,. \tag{13}$$

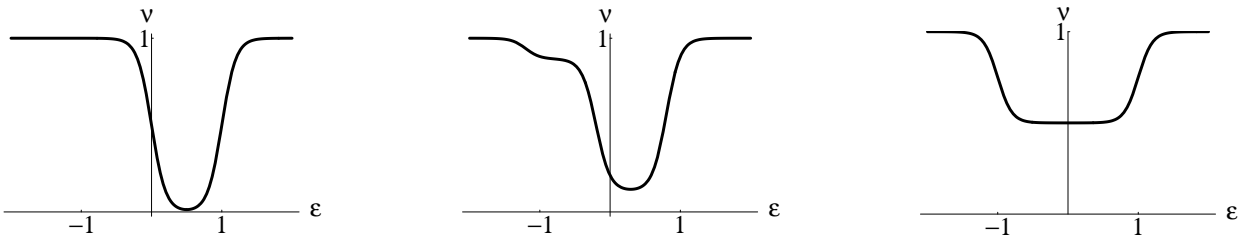

FIG. 2: The dependence of the TDoS (in the units of ν_0) on the energy (measured in E_c): (a) in the valley, (b) through an intermediate region, and (c) at the peak.

Although $G^<$ and $G^>$ are exactly related (in the equilibrium case) by $G_n^<(\varepsilon) = -\mathrm{e}^{\beta(\varepsilon-\mu)} G_n^>(\varepsilon)$, this is not convenient to use for approximations in the strong Coulomb blockade regime ($E_c\beta \gg 1$) as both functions in the product are very sharp. Instead, we note that there exists an exact relation, $Z_N = Z_N(\varepsilon_n) + e^{-\beta\varepsilon_n} Z_{N-1}(\varepsilon_n)$ that follows from the definition (9), which allows us to express $G^<$ after straightforward transformations as

$$G^<(\varepsilon) = \frac{2\pi i \nu_0}{\mathcal{Z}} \sum_N e^{-\beta E_N} f(\varepsilon - \Omega_{N-1}) \,. \tag{14}$$

From these expressions for the Green's functions we find the TDoS to be:

$$\frac{\nu(\varepsilon)}{\nu_0} = \frac{1}{\mathcal{Z}} \sum_N e^{-\beta E_N} \left[f\left(\varepsilon - E_c(N - \tfrac{1}{2} - N_g)\right) + 1 - f\left(\varepsilon - E_c(N + \tfrac{1}{2} - N_g)\right)\right] . \tag{15}$$

Keeping only the leading terms in the sum over N, we get

$$\frac{\nu(\varepsilon)}{\nu_0} = \frac{U(\varepsilon - \Omega_N) + e^{-\beta(\Omega_N - \mu)}\, U(\varepsilon - \Omega_{N+1})}{1 + e^{-\beta(\Omega_N - \mu)}}, \tag{16}$$

where we have defined $U(\varepsilon - \Omega_N) = f(\varepsilon - \Omega_{N-1}) + 1 - f(\varepsilon - \Omega_N)$. All other terms are exponentially suppressed so that we just have the terms for N closest to $N_g + \frac{1}{2}$. Away from the degeneracy point, one of the terms in eq. (16) is also exponentially suppressed so that this expression is contributed by one term only. This corresponds to the valley in the Coulomb blockade and gives the TDoS with a gap, as illustrated in fig. 2(a). In approaching the degeneracy point, fig. 2(b), the gap is smeared by the contribution from the other term in eq. (16); finally, at the degeneracy point corresponding to the peak of the Coulomb blockade, the TDoS remains finite at all energies but shows a 'half-gap' at $\varepsilon < E_c$, fig. 2(c).

Now we show that the half-gap in the TDoS is necessary to restore a correct expression for the current through an almost closed quantum dot. We consider such a dot connected via weak tunnelling contacts to leads (reservoirs) with fixed chemical potentials. Introducing in the usual way[16] the retarded, advanced and Keldysh components of the Green's function, one can rewrite the standard expression[17] for the current through the α^{th} lead (with $\Gamma_\alpha = 2\pi\nu_\alpha|\gamma_\alpha|^2$ being the tunnelling rate, and γ_α tunnelling coefficients) as follows:

$$\mathcal{I}_\alpha = e\Gamma_\alpha \int_{-\infty}^{+\infty} \frac{\mathrm{d}\varepsilon}{4\pi i} \,\mathrm{Tr}\left\{ \hat{G}^K(\varepsilon) - \left[1 - 2f_\alpha(\varepsilon)\right]\left[G^R(\varepsilon) - G^A(\varepsilon)\right]\right\} . \tag{17}$$

Restricting considerations to the case of two leads ($\alpha = 1, 2$) and using the current conservation, $\mathcal{I}_1 + \mathcal{I}_2 = 0$, we exclude the Keldysh compponent G^K from eq. (17) to find

$$\mathcal{I}_\alpha = \frac{e\Gamma_1\Gamma_2}{\Gamma} \int_{-\infty}^{+\infty} \frac{\mathrm{d}\varepsilon}{2\pi i} \left[f_2(\varepsilon) - f_1(\varepsilon)\right]\left[G^R(\varepsilon) - G^A(\varepsilon)\right] \tag{18}$$

In the linear response regime, when the difference between the chemical potentials in the leads $\mu_2 - \mu_1 = eV$, we obtain the following expression for the differential conductance,

$$\frac{\mathrm{d}\mathcal{I}}{\mathrm{d}V} = e^2\,\nu_0 \frac{\Gamma_1\Gamma_2}{\Gamma} \int_{-\infty}^{+\infty} \mathrm{d}\varepsilon \left[-\frac{\partial f(\varepsilon - \mu)}{\partial\varepsilon}\right] \frac{\nu(\varepsilon)}{\nu_0} \tag{19}$$

Here $\nu(\varepsilon)$ is the TDoS calculated for the dot in a thermodynamical contact with the leads at the same chemical potential μ. Substituting the expressions for the TDoS obtained above and keeping only the two leading order terms N and $N+1$ in the TDoS, i.e. using eq. (16), we arrive at

$$\int_{-\infty}^{+\infty} \mathrm{d}\varepsilon \left[-\frac{\partial f(\varepsilon-\mu)}{\partial \varepsilon} \right] \frac{\nu(\varepsilon)}{\nu_0} = \frac{\Omega_N/2T}{\sinh \Omega_N/2T},$$

provided that $|\Omega_N| = |E_a(N+\frac{1}{2}) - N_g| \ll E_a$. Substituting this into eq. (19) reproduces the standard result[1,2] for the differential conductance:

$$\frac{\mathrm{d}\mathcal{I}}{\mathrm{d}V} = e^2 \nu_0 \frac{\Gamma_1\Gamma_2}{\Gamma} \frac{\Omega_N/2T}{\sinh \Omega_N/2T}. \tag{20}$$

We note that for reproducing the correct coefficient at the peak in the differential conductance it is necessary to have the half-gap structure in the TDoS described above.

In conclusion, we have calculated the Green's functions describing an isolated quantum dot in the Coulomb blockade regime. This allows us to find a complete description of the TDoS both in the Coulomb valleys and peaks. At the peaks we find a new feature in the TDoS: the zero-energy suppression of the TDoS by the factor of 2 in comparison to its value at high energies. By considering the quantum dot weakly coupled to two leads we can find the linear response conductance for elastic tunnelling. We have also demonstrated that the correct description of the TDoS at the peak is necessary to obtain the standard result for the differential conductance. Finally, we believe that the techniques developed here could be extended for considerations of the spin blockade effects in the quantum dot[18].

Acknowledgments

This work was supported by the EPSRC grant GR/R95432.

[1] Alhassid Y., *Rev. Mod. Phys.* **72** (2000) 895.
[2] Aleiner I. L., Brouwer P. W. and Glazman L. I., *Phys. Rep.* **358** (2002) 309.
[3] Kulik I. O. and Shekhter R. I., *Zh. Eksp. Teor. Fiz.* **68** (1975) 623.
[4] Averin D. V. and Likharev K. K., *J. Low Temp. Phys.* **62** (1986) 345.
[5] Mullen K., Gefen Y. and Ben-Jacob E., *Physica B* **152** (1988) 172.
[6] Glazman L. I. and Matveev K. A. *JETP Lett.* **51** (1990) 484.
[7] Nazarov Y. V., *Zh. Eksp. Teor. Fiz.* **96** (1989) 975.
[8] Kamenev A. and Gefen Y., *Phys. Rev. B* **54** (1996) 5428.
[9] Levitov L. S. and Shytov A. V. *JETP Lett.* **66** (1997) 214.
[10] Altshuler B. L. and Aronov A. G., *Sol. State Commun.* **30** (1979) 115.
[11] Shklovskii B. I. and Efros A. L., *J. Phys. –Condens. Matter* **8**.
[12] Kamenev A. and Andreev A., *Phys. Rev.* B **60** (1999) 2218.
[13] Efetov K. B. and Tschersich A., *Europhys. Lett.* **59** (2002) 114; *Phys. Rev.* B **67** (2003) 174205.
[14] Kurland I. L., Aleiner I. L. and Altshuler B. L., *Phys. Rev.* B **62** (2000) 14886.
[15] Keldysh L. V., *Zh. Eksp. Teor. Fiz.* **47** (1964) 1515.
[16] Rammer J. and Smith H., *Rev. Mod. Phys.* **58** (1986) 323.
[17] Jauho A.-P., Wingreen N. S. and Meir Y., *Phys. Rev.* B **50** (1994) 5528.
[18] Kiselev M. N. and Gefen Y., *Phys. Rev. Lett.* **96** (2006) 066805.